

VỊ TRÍ CỦA TIẾNG CAO LAN TRONG CÁC NGÔN NGỮ TAI

ThS PHAN LƯƠNG HÙNG

1. Người Cao Lan và tiếng Cao Lan

Theo danh mục các dân tộc ở Việt Nam do Tổng cục Thống kê ban hành năm 1979, người Cao Lan là một nhóm địa phương thuộc dân tộc Sán Chay. Kết quả điều tra dân số năm 2009 cho thấy có 169.410 người Sán Chay, sống tập trung tại các tỉnh trung du miền núi phía Bắc: Tuyên Quang (61.343 người), Thái Nguyên (32.483 người), Bắc Giang (25.821 người), Quảng Ninh (13.786 người), Yên Bái (8.461 người), Cao Bằng (7.058 người), Lạng Sơn (4.384 người), Phú Thọ (3.294 người), Vĩnh Phúc (1.611 người), Bắc Kạn (602 người)... Một số mới di cư vào các tỉnh miền Trung, Tây Nguyên và Nam Trung Bộ.

Dân tộc Sán Chay có hai nhóm địa phương là Cao Lan và Sán Chỉ. Trong đó, người Cao Lan chiếm khoảng 63% tổng dân số dân tộc Sán Chay. Người Cao Lan tự gọi mình là Hồn Bán, có nghĩa là "người ở bản". Ở Thái Nguyên, Tuyên Quang và Yên Bái cộng đồng này còn tự gọi mình là Sán Cháy. Ngoài ra, họ còn được biết đến với những tên gọi khác như: Sùn Nhần (người ở thôn bản), Phén, Chùng...

Người Sán Chỉ tự gọi mình là Sán Chay, có nghĩa là "quả ở trên rừng". Cũng có ý kiến cho rằng tộc danh Sán Chay bắt nguồn từ hai chữ *Son Tử*, nghĩa là người ở trên núi [4, 55].

Theo một số tài liệu dân tộc học, người Cao Lan vốn cư trú ở khu vực Nam Trung Quốc. Sau đó, do chiến tranh, loạn lạc, dịch bệnh... họ di cư đến khu vực Quảng Đông, Quảng Tây Trung Quốc rồi vào Việt Nam bằng nhiều con đường khác nhau. Hiện nay, các bài cúng của người Cao Lan còn nhắc tới Quảng Đông, Quảng Tây như là quê hương của họ. Các gia phả hiện còn lưu giữ được trong cộng đồng người Cao Lan cũng cho thấy thời gian họ có mặt ở Việt Nam chỉ mới khoảng 400 năm nay [4, 66].

Tiếng Cao Lan là một ngôn ngữ đơn tiết, có thanh điệu. Về phân loại thân tộc ngôn ngữ, hiện nay tất cả các nhà nghiên cứu đều mới chỉ thống nhất với nhau về vị trí của tiếng Cao Lan là thuộc ngữ hệ Tai - Kadai, nhánh Kam - Tai, tiểu nhánh Be - Tai, nhóm Tai. Vấn đề tiếng Cao Lan thuộc tiểu nhóm nào trong nhóm Tai vẫn còn chưa được thống nhất.

Ngôn ngữ là công cụ của tư duy và là công cụ giao tiếp chính của mỗi cộng đồng. Nó chứa đựng nhiều trầm tích về văn hóa và cả lịch sử của một tộc người. Hiện nay, ngôn ngữ được coi là một trong những tiêu chí quan trọng nhất đối với việc xác định thành phần dân tộc (bên cạnh các tiêu chí khác như ý thức tự giác tộc người hay các đặc trưng văn hóa). Bài viết này của chúng tôi sẽ áp dụng thủ pháp thống kê từ vựng (lexico-statistic) và xác định những cách tân ngữ âm (innovation) để xác định vị trí của tiếng Cao Lan trong các ngôn ngữ Tai. Những kết quả của bài viết này cũng sẽ góp phần xác định thành phần tộc người, đồng thời làm rõ hơn quá trình lịch sử của cộng đồng Cao Lan.

Tư liệu được sử dụng trong bài viết bao gồm bảng 200 từ của Swadesh¹ được phiên âm quốc tế tiếng Cao Lan và 03 ngôn ngữ đại diện cho các tiểu nhóm Tai khác (tiếng Giáy, tiếng Tày và tiếng Thái) do chúng tôi trực tiếp thu thập tại Tuyên Quang (2011), Bắc Giang (2004), Hòa Bình (2006) và kế thừa kết quả nghiên cứu của các công trình đi trước.

2. Lịch sử nghiên cứu tiếng Cao Lan

Trên bình diện đồng đại, tiếng Cao Lan hầu như chưa được miêu tả một cách có hệ thống và kỹ lưỡng. Công trình *Ngữ âm tiếng Cao Lan* của Phan Lương Hùng (2010) mới chỉ dừng lại ở những mô tả bước đầu về hệ thống ngữ âm của ngôn ngữ này. Trong khi đó, các công trình nghiên cứu về tiếng

Cao Lan trên bình diện lịch đại lại khá nhiều, thí dụ như: *The classification of the Cao Lan language* của David Strecker, *Những vấn đề âm vị học lịch đại* của Haudricourt, *Some puzzles in Cao Lan* của Gregerson và Edmondson... Nhìn chung, các nhà nghiên cứu có 03 quan điểm khác nhau về vị trí của tiếng Cao Lan trong các ngôn ngữ Tai như sau:

Thứ nhất, tiếng Cao Lan thuộc nhóm Tai Bắc. Tiêu biểu cho các học giả ủng hộ quan điểm này là J.Gregerson và J.Edmonson [13], David Strecker [11], Nguyễn Văn Lợi [10]...

Thứ hai, tiếng Cao Lan thuộc nhóm Tai Trung tâm. Đây là quan điểm được các học giả Trung Quốc ủng hộ: “tiếng Cao Lan không phải là tiếng Thái, mà là một phương ngữ Choang của huyện Khâm... Các phương ngữ Choang ở miền Đông Nam khác với các phương ngữ Choang ở các vùng khác ở chỗ có một hệ thống phụ âm gần với tiếng Tày - Nùng” [3, 5]. Đây cũng là quan điểm chính thức được đăng tải trên trang web của Viện Ngôn ngữ học Mùa hè (Summer Institute of Linguistics) của Mỹ về các ngôn ngữ trên thế giới [12].

Thứ ba, tiếng Cao Lan không thuộc Tai Bắc hay Tai Trung tâm. Nó cùng với Nùng An và T'sun Va tạo thành một nhóm riêng trong các ngôn ngữ Tai. Nhóm này được gọi là nhóm Cao Lan. Đây là quan điểm chính thức được A.G.Haudicourt nêu lên trong bài viết *Note sur les dialectes de la region de Moncay* công bố năm 1960.

Trong các công trình nêu trên, nhìn chung các tác giả mới chỉ dừng lại ở một vài hiện tượng đơn lẻ [7a], [6], [11], hoặc không đưa ra những chi tiết cụ thể khi bàn về vị trí của tiếng Cao Lan [9].

3. Vị trí của tiếng Cao Lan trong các ngôn ngữ Tai

Để xác định vị trí của một ngôn ngữ trong cây phả hệ, hay nói cách khác là minh định mức độ thân thuộc giữa các ngôn ngữ có quan hệ cội nguồn, các nhà nghiên cứu thường dùng thủ pháp so sánh từ vựng và tìm các xu hướng cách tân từ tiền ngôn ngữ cho đến các ngôn ngữ hiện đại. Đối với thủ pháp so sánh từ vựng, chúng tôi sử dụng bảng 100 và 200 từ của Morris Swadesh. Đây là danh mục các từ vựng cơ bản nhất của một ngôn ngữ được Swadesh đưa ra trong khi phát triển phương pháp ngữ thời học (glottochronology) để xác định thời gian chia tách giữa các ngôn ngữ. Theo đó, ông giả định rằng cũng như sự phân rã các-bon đối với các loại vật chất thì theo thời gian, từ vựng của một ngôn ngữ cũng có sự phân rã nhất định. Trong các lớp từ vựng thì lớp từ cơ bản có tốc độ phân rã ổn định nhất và cũng ít bị vay mượn nhất. Do vậy, việc sử dụng các từ cơ

bản sẽ cho chúng ta kết quả chính xác nhất có thể về thời điểm chia tách của các ngôn ngữ có mối quan hệ thân tộc cũng như mức độ gần gũi giữa chúng.

Sự chia tách ngôn ngữ, sau đó cũng sẽ tiếp tục dẫn tới những cách tân (innovation) hay bảo lưu (retention) khác nhau do chúng đã chịu những tác nhân bên trong hay bên ngoài ngôn ngữ khác nhau. Nói cách khác là kể từ sau thời điểm chia tách, mỗi ngôn ngữ được chia tách sẽ trải qua những quá trình phát triển riêng. Cách tân ngôn ngữ diễn ra sau thời điểm chia tách ngôn ngữ và do vậy, chúng thường mang những đặc điểm riêng. Vậy, những ngôn ngữ bảo lưu cùng một hay một số đặc điểm cách tân chắc chắn đã có cùng một ngôn ngữ mẹ sau thời điểm chia tách [3, 167].

Với những tiền đề nêu trên, chúng tôi đã tiến hành so sánh 200 từ vựng cơ bản theo bảng từ của Morris Swadesh giữa tiếng Cao Lan với tiếng Tày (đại diện tiểu nhóm Tai Trung tâm), tiếng Thái (đại diện tiểu nhóm Tai Tây Nam) và tiếng Giáy (đại diện tiểu nhóm Tai Bắc). Dưới đây là bảng kết quả so sánh từ vựng:

Ngôn ngữ so sánh	Giáy	Thái	Tày
Cao Lan	60,11%	54,12%	46,94%

Bảng trên cho thấy rõ ràng là tiếng Cao Lan và các ngôn ngữ Tai có mối quan hệ cội nguồn như các nhà nghiên cứu đã khẳng định. Điều này được thể

hiện qua tỉ lệ từ vựng tương đồng cao giữa các ngôn ngữ được so sánh ở trên và các từ chung này được trải đều, một cách có hệ thống trong các trường từ vựng như hiện tượng thiên nhiên: *mặt*

trời, sao, nước, đất, đá; bộ phận cơ thể: đầu, tóc, mắt, mũi, răng; đại từ nhân xưng và từ chỉ quan hệ thân tộc: con, cháu, tao, mày; từ chỉ trạng thái:

nhiều, dài, xanh, vàng, sưng... Dưới đây là một số thí dụ về từ chung giữa tiếng Cao Lan với tiếng Tày, tiếng Thái và tiếng Giáy:

Việt	Cao Lan	Tày	Thái đen	Giáy
*mặt trôi	t ^h āk ³⁵ Nen ³²	pHjA ⁴⁵ vA(n ³³)	ta ³³ ven ³³	tAN ⁴⁴ va&n ^{55/}
*sao	daw ⁴⁴ dΦj ³²	dAw ⁴⁵ di ³³	dao ³³	daw ⁴⁴ di ²¹
* [®] Êt	tom ⁴⁴	din ⁴⁵	din ³³	nam ^{55/}
*nước	nΦ(m ^{3...43})	nA(m ²¹)	n∅m ⁵⁵	Δa&m ⁴⁵
* [®] ,	then ⁴⁴	hin ⁴⁵	hin ³⁵	Δin ⁴⁴
*tao	kΦ(w ⁴⁴)	kA(w ³²³)	ku ³³	ku ⁴⁴
*muy	mΦ(N ⁴⁴)	mΦ(∞ ³³)	m∞N ³³	m∞N ⁵⁵
*bụng	tuN ^{3...43}	pAN ³²³	p∞m ³³	tuN ⁴⁵
*gan	t [”] p ³⁵	tA(p ⁴⁵)	t∅p ³⁵	ta&p ⁴⁴
*xanh	hEw ⁴⁴	kHEw ³²³	ΞE□ ³⁵	l□k ²²
*vung	l∞N ⁴⁴	l∞ΦN ³²³	l∞↔N ³⁵	hEn ³²⁵
*đui	l [”] j ³²	ŕi ⁵⁵	hi ³³	Δa&φ ⁵⁵
*nhiều	laj ⁴⁴	lAj ³²³	laj ³⁵	laj ⁴⁴

Chúng ta thấy tỉ lệ tương đồng về mặt từ vựng giữa tiếng Cao Lan với tiếng Giáy là lớn nhất (60,11%); sau đó đến tiếng Thái (54,12%), và cuối cùng là tiếng Tày (46,94%). Nói cách khác, về mặt từ vựng thì tiếng Cao Lan gần với tiếng Giáy nhất. Để có thêm cơ sở phân tích, chúng tôi cũng tiến hành thống kê so sánh từ vựng cơ bản chung giữa ba ngôn ngữ

Tày, Thái và Giáy vốn thuộc ba tiểu nhóm Tai khác nhau. Kết quả cho thấy tỉ lệ từ cơ bản chung giữa ba ngôn ngữ này cao nhất là 51% (Thái - Giáy), tiếp đó lần lượt giảm dần xuống 47,67% (Thái - Tày) và 44,15% (Tày - Giáy). Các tỉ lệ này nhìn chung tương đương với tỉ lệ từ chung Cao Lan - Tày và Cao Lan - Thái. Nói cách khác, thống

kê từ vựng cơ bản Cao Lan - Tày và Cao Lan - Thái cho ta kết quả của các ngôn ngữ thuộc các tiểu nhóm Tai khác nhau.

Riêng trường hợp tỉ lệ từ cơ bản chung Cao Lan - Giáy cao bất thường (60,11%) so với các trường hợp còn lại khiến chúng ta có thể nghĩ đến khả năng tiếng Cao Lan và tiếng Giáy cùng thuộc một tiểu nhóm - tiểu nhóm Tai Bắc. Tỉ lệ 60,11% tương đồng về từ vựng có đủ để chúng ta khẳng định được rằng tiếng Cao Lan và tiếng Giáy cùng thuộc một tiểu nhóm - tiểu nhóm Tai Bắc hay không? Về mặt phương pháp, chúng ta còn phải dựa vào những cứ liệu về cách tân ngữ âm. Tuy nhiên, ngay trong lĩnh vực từ vựng, bên cạnh tỉ lệ tương đồng ra chúng ta cũng có thể tham khảo những từ chứng do Lý Phương Quế [10a] đưa ra. Trong công trình của mình, ông đã dựa trên những quan sát thực tế về từ vựng trong các

ngôn ngữ Tai và đưa ra danh mục từ chứng đối với các tiểu nhóm Tai. Ông cho rằng: "...không thể nói rằng sự phân bố từ vựng là chìa khóa quan trọng nhất hay duy nhất trong việc phân chia ngôn ngữ, nhưng rõ ràng là từ vựng là nhân tố cơ bản cả trong phân loại cội nguồn (genetic classification) và phân loại địa lí ngôn ngữ/ phương ngữ (dialect geography). Vậy nếu phân loại ngôn ngữ dựa trên từ vựng và được hỗ trợ bởi tiêu chí âm vị học thì sẽ có độ tin cậy (validity) cao hơn rất nhiều..." [10a, 16]. Tuy quan điểm lấy từ vựng làm trung tâm trong phân loại cội nguồn của ông còn nhiều tranh cãi và sau này đã bị chính ông từ bỏ trong công trình [10b] nhưng nó cũng cho thấy tầm quan trọng của từ vựng trong xác định mối quan hệ cội nguồn. Đối chiếu với các từ chứng mà Lý Phương Quế đã đưa ra, chúng tôi thu được kết quả như sau:

Từ được ss Ngôn ngữ ss	<i>hố</i>	<i>làm</i>	<i>đếm</i>	<i>lạnh</i>	<i>mây</i>	<i>lười</i>
Cao Lan	kok ^{32□}	kuk ³¹	ʁun ⁴⁴	cep ³⁵	pH ³ ₄	kec ^{32□}
Giáy	ku#k ²²	ku ⁵⁵	/an ²¹	ca&t ⁴⁴	v ³² Φ ³² ₅	ci&k ²²
Thái	s ³⁵ Φ ³⁵ ₄	zet ⁴²	n̄p ³¹	naw ³⁵⁴	fa ^{33/}	can ³³
Lự	sΦ ³³	het ³¹	na&p ³ ₁	naw ³³	fa ¹³	Ξan ¹¹
Tày	▷ ⁵ Φ ⁵ ₅	hE(t ¹¹	/an ³³	jen ³³	pHa ³⁵	zan ³²³

(ss: So sánh)

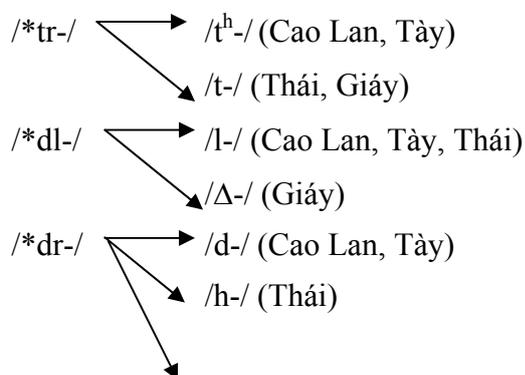
Lý Phương Quế cho rằng đối với các từ *hổ, làm, mây* và *lười*, các ngôn ngữ Tai Bắc (mà ở đây đại diện là tiếng Giáy) có dạng thức riêng biệt, khác biệt hoàn toàn với các ngôn ngữ Tai khác. Bảng trên cho ta thấy rõ dạng thức của các từ chứng này trong tiếng Cao Lan rất gần gũi với tiếng Giáy. Như vậy, bên cạnh những đặc trưng của các ngôn ngữ Tai nói chung thì

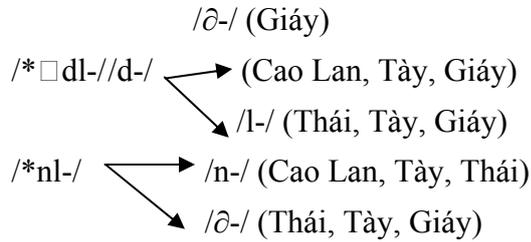
từ vựng tiếng Cao Lan mang đậm đặc trưng của các ngôn ngữ tiểu nhóm Tai Bắc.

Về xu hướng cách tân từ Proto-Tai đến tiếng Cao Lan, chúng ta hãy xem xét một số trường hợp biến đổi của loạt tổ hợp phụ âm răng từ Proto-Tai sang tiếng Cao Lan, Tày, Thái và Giáy sau đây:

Việt	Cao Lan	Tày	Thái đen	Giáy	Proto-Tai
<i>mất</i>	t ^h a ³³	tHa ⁵⁵	ta ³²	ta ⁴⁴	*tr
<i>chét</i>	tHaj ³³	tHaj ⁵⁵	taj ³²	taj ⁴⁴	*tr
<i>đuôi</i>	t ^h ∞N ³³	tHaN ⁵⁵	haN ³²	T∞ ^l ∞N ⁴⁴	*tr
<i>gió</i>	lom ³³	lo&m ³³	lom ³³	Δum ^{55/}	*dl
<i>xuống</i>	loN ³²	loN ³³	loN ³³	Δ□N ⁵⁵	*dl
<i>rẽ</i>	lak ³²	lak ³³	ha ³¹	∂ak ³⁵	*dr
<i>(ngày) kia</i>	l□n ³²	l∞ ³³	h∞ ³⁵	∂∞ ⁴⁵	*dr
<i>sao</i>	daw ³³	daw ³³	law ³²	daw ⁴⁴	* □ dl
<i>núi</i>	d□j ³³	pu ³³	pu ³³	po ⁴⁴	* □ dl
<i>nước</i>	num ³²³	na&m ³²³	năm ²⁴³	∂a&m ⁴⁵	*nl
<i>chim</i>	nok ^{32□}	no&k ³²³	nok ³¹	∂□k ²²	*nl

Bảng trên cho ta thấy các xu hướng cách tân như sau:





Chúng ta thấy xu hướng cách tân trong tiếng Cao Lan gần nhất với xu hướng cách tân trong tiếng Tày ở cả 5/ 5 trường hợp. Tiếp đó đến tiếng Thái với 2/ 5 trường hợp và cuối cùng là tiếng Giáy chỉ với 1/ 5 trường hợp. Như vậy, xét về phương diện xu hướng

cách tân ngữ âm thì tiếng Cao Lan lại gần nhất với tiếng Tày - một ngôn ngữ Thái Trung tâm. Ta có thể tổng hợp cả hai tiêu chí về mức độ gần gũi từ vựng và cách tân ngữ âm giữa tiếng Cao Lan với 3 ngôn ngữ còn lại trong bảng sau:

Ngôn ngữ ss \ Tiêu chí ss	Tày	Thái	Giáy
Từ vựng	1	2	3
cách tân ngữ âm	3	2	1

(1- mức độ gần gũi thấp nhất, 3- mức độ gần gũi cao nhất)

Bảng trên cho thấy xu hướng trái ngược nhau về mức độ gần gũi giữa tiếng Cao Lan với tiếng Tày và tiếng Giáy ở hai tiêu chí từ vựng và xu hướng cách tân ngữ âm. Nếu như tiếng Cao Lan gần với tiếng Giáy nhất và xa tiếng Tày nhất về từ vựng thì ngược lại, nó lại gần tiếng Tày nhất và xa tiếng Giáy nhất về cách tân ngữ âm.

Đứng trước tình huống này, nếu lấy tiêu chí từ vựng làm trọng tâm thì ta có thể xếp tiếng Cao Lan vào cùng nhóm với tiếng Giáy, tức là cùng thuộc tiểu nhóm Tai Bắc. Nếu lấy tiêu chí xu hướng cách tân làm trọng tâm thì ta phải xếp tiếng Cao Lan vào cùng nhóm với tiếng Tày, tức là cùng thuộc tiểu nhóm Tai Trung tâm.

4. Một vài nhận định

Như chúng tôi đã nêu ở trên, các tài liệu dân tộc học đều khẳng định rằng người Cao Lan vốn từng cư trú ở vùng Quảng Đông, Quảng Tây (Trung Quốc), đã di cư sang Việt Nam vào nhiều thời điểm khác nhau và bằng nhiều con đường khác nhau. Cuốn *Giới thiệu các dân tộc thiểu số ở Tuyên Quang*, 1972 cũng đã viết rõ “Sách hương hỏa của đồng bào thường ghi: Đại Minh quốc, Quảng Đông tỉnh, Hải Bắc đạo, Liêm châu sở, Khâm châu huyện...” [1, 90]. Chúng ta cũng đều biết rằng Quảng Đông và Quảng Tây vốn là khu vực cư trú lâu đời của cư dân Bách Việt (Bai Yue). Tên gọi Choang (Zhuang) chỉ chính thức xuất hiện từ thời Tống và được dùng cho đến ngày nay. Cộng đồng Choang ở Trung Quốc

được chia làm hai nhóm: Choang Bắc và Choang Nam. Đúng như tên gọi của chúng, Choang Bắc nói một ngôn ngữ thuộc tiểu nhóm Tai Bắc; trong khi Choang Nam lại nói một ngôn ngữ thuộc tiểu nhóm Tai Trung tâm. Về đại thể, Choang Bắc có xu hướng tập trung hơn về phía Bắc vùng Lưỡng Quảng trong khi Choang Nam lại tập trung đông hơn ở khu vực phía Nam. Tuy nhiên, nhìn chung không có một ranh giới rõ ràng và tình trạng cư trú đan xen là phổ biến. Như vậy, ta có thể hình dung các khả năng sau về nguồn gốc tộc người của cộng đồng Cao Lan:

Thứ nhất, người Cao Lan vốn là cư dân Choang Bắc, sau đó do tiếp xúc cộng cư với các cư dân Choang Nam lâu đời nên đã bị ảnh hưởng bởi cách phát âm của tiếng Choang Nam.

Thứ hai, người Cao Lan vốn là cư dân Choang Nam, sau đó cũng do tiếp xúc cộng cư mà vay mượn nhiều từ vựng tiếng Choang Bắc. Trong đó có cả những từ đặc trưng cho ngôn ngữ Choang Bắc.

Thứ ba, người Cao Lan vốn là cư dân Choang nói một phương ngữ Dao². Sau đó, trong quá trình cộng cư lâu đời với cộng đồng nói tiếng Tai đông hơn nên đã chuyển sang nói một ngôn ngữ Tai như Haudricourt [7c] đã giả định.

Chúng tôi lựa chọn giả thuyết thứ nhất bởi vì trên thực tế, từ vựng cơ bản là yếu tố có “độ phân rã” thấp nhất, chậm nhất trong toàn bộ từ vựng của một ngôn ngữ. Ở trên chúng ta cũng đã thấy tỉ lệ từ chung cùng nguồn gốc cao giữa tiếng Cao Lan với tiếng Giáy và sự có mặt đều đặn của các từ

chứng của tiểu nhóm Tai Bắc trong tiếng Cao Lan. Trong khi đó, ngữ âm lại là yếu tố dễ và sớm bị ảnh hưởng hơn theo thời gian do những nhân tố bên trong và ngoài ngôn ngữ.

Trong đợt điền dã tại Bắc Giang (2004), chúng tôi đã được cụ Tổng Văn Bình, 76 tuổi, thôn Công Luộc, xã Đèo Gia cho biết: "trước đây, Cao Lan như người vợ lẽ trong gia đình, chỉ được sống ở dưới nhà ngang". Có một thực tế là ngôn ngữ của cộng đồng có vị thế thấp thường chịu ảnh hưởng của ngôn ngữ của cộng đồng có vị thế cao hơn. Người Cao Lan rất có thể vốn là cư dân Choang Bắc, di cư xuống phía Nam vùng Lưỡng Quảng và cư trú ở khu vực có đa số cư dân nói tiếng Choang Nam. Trong trường hợp này, cách phát âm của người Cao Lan bị ảnh hưởng là điều dễ hiểu. Sự có mặt của từ vựng gốc Dao như: *pam*³² (bùn)... là do quá trình tiếp xúc lâu đời và nó chỉ lưu lại dấu ấn tiếp xúc này ở một vài từ lẻ tẻ.

Tóm lại, từ tất cả những cứ liệu và phân tích ở trên, chúng tôi cho rằng tiếng Cao Lan vốn thuộc tiểu nhóm Tai Bắc. Trong quá trình cư trú và di cư, họ đã tiếp xúc với các cộng đồng nói các ngôn ngữ Tai Trung tâm và lưu lại dấu ấn của các xúc tiếp này trong hệ thống ngữ âm của mình. Nguồn gốc Tai Bắc của tiếng Cao Lan vẫn còn được lưu giữ trong vốn từ vựng cơ bản của họ. Đặc biệt là các từ chứng Tai Bắc vẫn được bảo lưu.

Chúng tôi vẫn cho rằng để có kết luận thuyết phục hơn về vấn đề này cần phải có một công trình nghiên cứu sâu hơn, xem xét kĩ hơn cả về từ vựng

lẫn các quá trình cách tân ngữ âm trên một diện tư liệu lớn hơn ở nhiều ngôn ngữ Tai khác nhau. Đây cũng là nhiệm vụ được chúng tôi đặt ra trong các nghiên cứu tiếp sau.

CHÚ THÍCH

¹ Nhà ngôn ngữ học người Mỹ - cha đẻ của phương pháp ngữ thời học.

² Tên gọi Choang cũng được dùng để chỉ chung cộng đồng dân tộc thiểu số sống ở Quảng Đông và Quảng Tây nói các ngôn ngữ Tai, Yao và Mông.

TÀI LIỆU THAM KHẢO

1. Ban Dân tộc tỉnh Tuyên Quang, *Giới thiệu sơ lược các dân tộc thiểu số ở Tuyên Quang* (Lưu hành nội bộ), 1972.
2. August Bonifacy, *Étude sur les Cao Lan*, T'oung Pao ou Archives concernant L'histoire, Les langues, La géographie et l'ethnographie de l'Asie orientale, Serie ii, Vol.viii, Leiden: E.J Brill, Pg429-438, 1907.
3. Terry Crowley, *An introduction to historical linguistics*, Oxford University Press, 1992.
4. Khổng Diễn, *Dân tộc Sán Chay ở Việt Nam*, Nxb Văn hóa Dân tộc, H., 2002.
5. Jerold A. Edmondson and David B. Solnit, *Comparative Kadai - The Tai Branch*, University of Texas at Arlington, 1997.
6. Kenneth J. Gregerson and Jerold A. Edmondson, *Some puzzles in Cao Lan*, Proceedings of the International Conference on Tai Studies, Mahidol University, Bangkok, pp151-63, 1998.
7. Andre G. Haudricourt
 - a. *Notes sur les dialectes de la région de Moncay*, BEFFEO 50, pg167-177, 1960.
 - b. *Những vấn đề âm vị học lịch đại*, Tập 1, Bản dịch Viện ngôn ngữ học, H., 1972.
 - c. *Mấy điều nhận xét về lí luận và thực tiễn nhận một chuyến đi thăm các dân tộc thiểu số ở Tây Bắc và Việt Bắc Việt Nam*, T/c Ngôn ngữ, Số 3, 1973.
8. Phan Lương Hùng
 - a. *Thanh điệu tiếng Thái Vạn Mai*, T/c Ngôn ngữ và Đời sống, Số 9, 2007.
 - b. *Ngữ âm tiếng Cao Lan*, Luận văn Thạc sĩ, ĐH KHXH&NV, 2010.
9. Nguyễn Văn Lợi, *Quan hệ Cao Lan - Sán Chí xét về mặt ngôn ngữ*, Hội thảo Xác định thành phần dân tộc Sán Chay (Cao Lan - Sán Chí), Viện Ngôn ngữ học - Viện Dân tộc học, Bắc Giang, 2004.
10. Lý Phương Quế,
 - a. *Classification by vocabulary: Tai dialects Anthropological Linguistics 1*, 1959.
 - b. *A Hand book of comparative Tai*, The Univ, Press of Hawaii, USA, 1977.
11. David Strecker, *The classification of the Cao Lan languages*, In S. Ratanakul et al eds, *Southeast Asian Linguistic Studies Presented to Andre-G. Haudricourt*, Mahidol University, Bangkok, tr.479-492, 1985.
12. Summer Institute of Linguistics, www.ethnologue.com.

13. Morris Swadesh, *Định niên đại tuyệt đối các xúc tiếp tộc người thời tiền sử bằng phương pháp thống kê từ vựng* (bản dịch Viện Ngôn ngữ học).

14. Graham Thurgood, *Các ngôn ngữ tai Kadai và Nam Đảo - Bản chất của mối quan hệ lịch sử*, T/c Ngôn ngữ, Số 3, 1998.

SUMMARY

The lexicostatistical method has been used in this article to clarify the position of Cao Lan in Tai languages. Swadesh's wordlist includes 200 basic words was compared between Cao Lan and three other Tai languages representing for Central Tai (Tay language), Southwestern Tai (Thai language) and Northern Tai (Giay language). Results of this statistic as well as testing-words characterizing subgroups of Tai languages lead us to the conclusion that Cao Lan language belongs to Northern Tai.